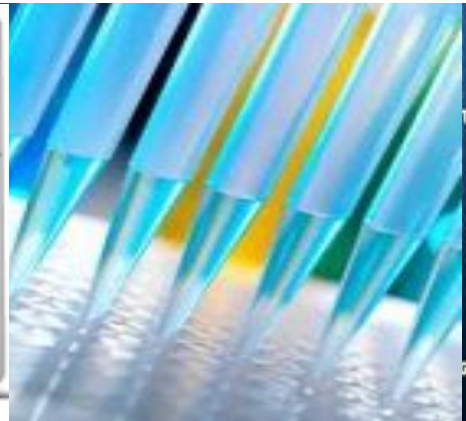
NGS for NBS

Are we ready for « Next Generation » Newborn Screening ?

François BOEMER
Biochemical Genetics Laboratory
Metabolics.be – 12/05/2017





Background

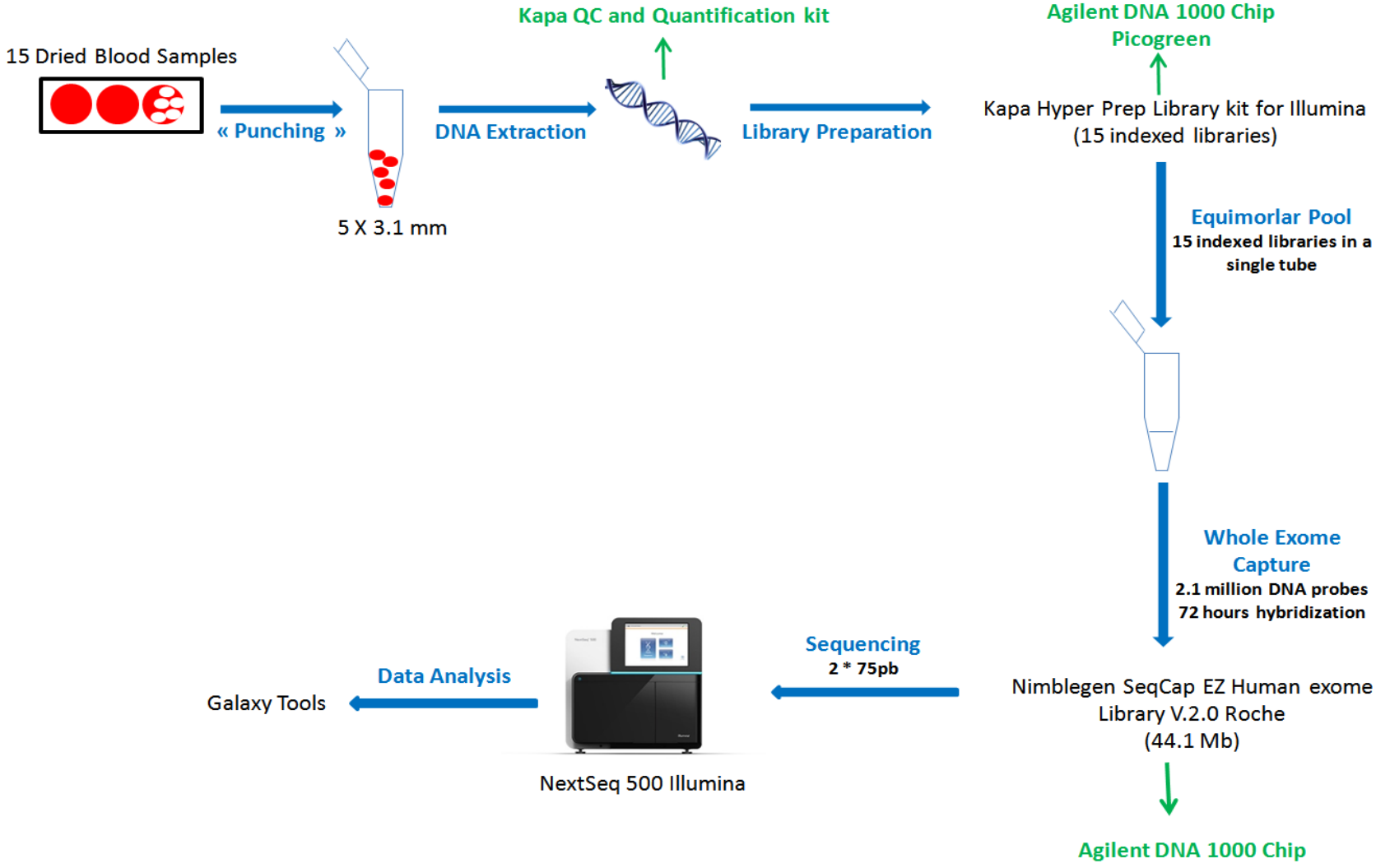
- NGS
  - Millions of genomes, exomes, transcriptomes over the last decades
  - Validated mainly on high quality DNA or FFPE samples, some papers on DBS reporting only technical capacities
- Filter paper
  - Evolved as a reference procedure for collection, transport, analysis and storage of biological fluids (ie CLSI guidelines)
  - Applications : NBS, diet follow-up, TDM and doping control, viral load measurements, targeted genes sequencing,…
- Objectives
  - Evaluate WES on DBS material – 15 IEM patients
    - Quality and quantity of retrieved DNA
    - Detection of known mutation
    - Focus on NBS genes to demonstrate suitability for future NBS programs

Patients

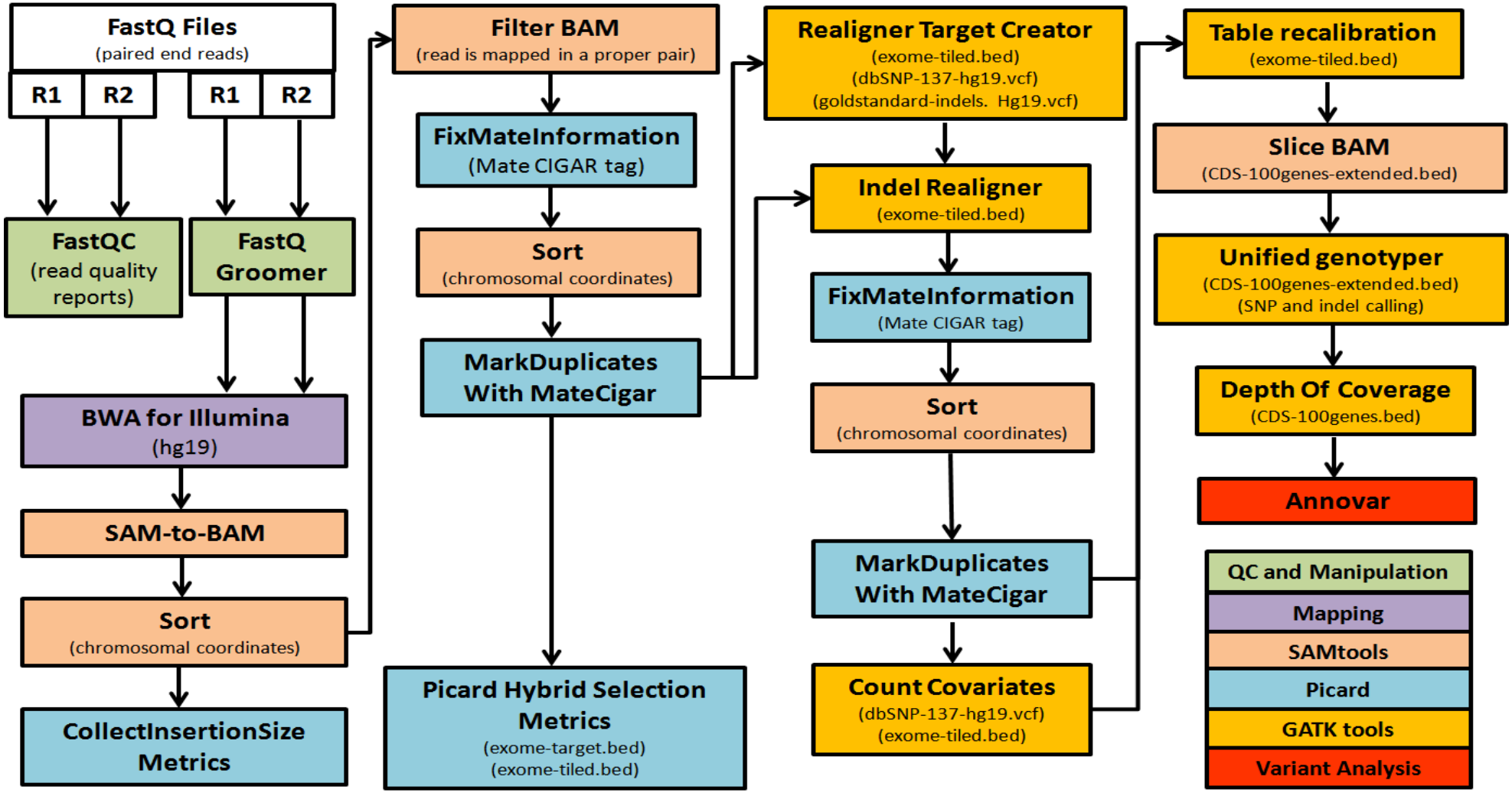
Patient's Id	Disorder	Gene	Mutation(s) Allele 1	Mutation(s) Allele 2	Comment
DBS-1	PKU	*PAH*	c.482T>C	c.1222>T	
DBS-2	PKU	*PAH*	c.473G>A	c.842C>T	
DBS-3	MCAD	*ACADM*	c.985A>G	c.985A>G	
DBS-4	Propionic Acidemia	*PCCB*	c.990_991 insT	c.1252G>A	
DBS-5	Methylmalonic Aciduria	*MMAB*	c.556C>T	c.563-577dup	
DBS-6	Tyrosinemia type I	*FAH*	c.554-1G>T	c.554-1G>T	
DBS-7	Glutaric Aciduria type I	*GCDH*	c.371G>A	c.1204C>T	
DBS-8	3-MCC[a]	*MCCC2*	c.1423G>A	c.1535A>C	
DBS-9	Propionic Acidemia	*PCCB*	c.997delA	c.763G>A	
DBS-10	Homocystinuria	*CBS*	c.429C>G	c.833T>C	
DBS-11	PKU	*DHPR*	c.661C>T	c.661C>T	
DBS-12	Galactosemia	*GALT*	c.563A>G	c.-119delGTCA c.378-27G>C c.508-24G>A c.507+62G>A c.940A>G	Allele 2 corresponds to Duarte 2 haplotype
DBS-13	MADD	*ETDFH*	c.293T>A	c.293T>A	
DBS-14	MSUD[b]	?[c]	?[c]	?[c]	
DBS-15	MCAD	*ACADM*	c.985A>G	c.1091T>C	

[a] Disorder not mandated by the newborn screening program of French community of Belgium.
[b] Maple Syrup Urine Disease
[c] Sanger sequencing has not been performed for MSUD patient.

Analytical workflow



Framework for variation discovery



Studied genes

- 35 IEM genes considered by NBS program in Wallonia
- 65 genes considered by diverse NBS programs worldwide

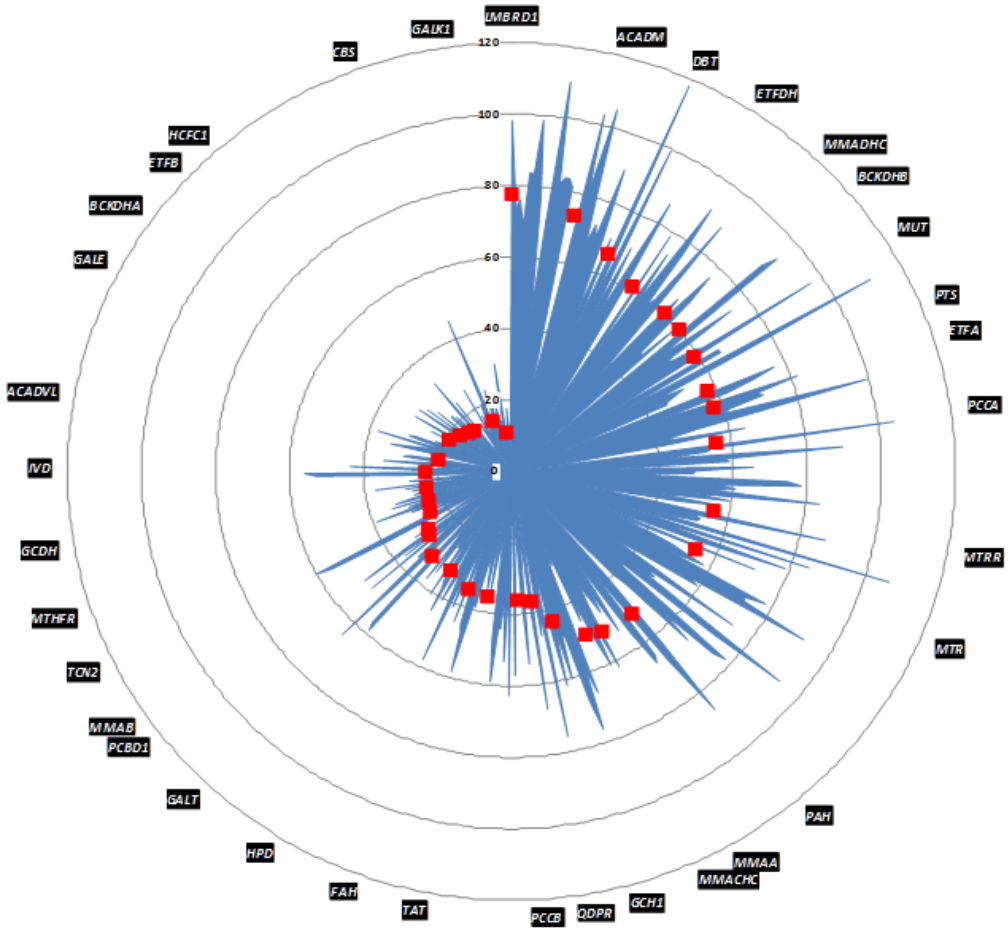
A. IEM currently screened in FWB	Genes
Phenylketonuria	*PAH*
Phenylketonuria	*PTS*
Phenylketonuria	*GCH1*
Phenylketonuria	*QDPR*
Phenylketonuria	*PCBD1*
MSUD	*DBT*
MSUD	*BCKDHA*
MSUD	*BCKDHB*
Tyrosinemia	*FAH*
Tyrosinemia	*TAT*
Tyrosinemia	*HPD*
Homocystinuria	*CBS*
Homocystinuria	*MTHFR*
Homocystinuria	*MTRR*
Homocystinuria	*MTR*
Galactosemia	*GALT*
Galactosemia	*GALK1*
Galactosemia	*GALE*
Methylmalonic Acidemia	*MUT*
Methylmalonic Acidemia	*MMACHC*
Methylmalonic Acidemia	*MMADHC*
Methylmalonic Acidemia	*LMBRD1*
Methylmalonic Acidemia	*HCFC1*
Methylmalonic Acidemia	*MMAA*
Methylmalonic Acidemia	*MMAB*
Methylmalonic Acidemia	*TCN2*
Propionic Acidemia	*PCCA*
Propionic Acidemia	*PCCB*
Glutaric Aciduria type I	*GCDH*
Isovaleric Acidemia	*IVD*
MCAD	*ACADM*
MADD	*ETFDH*
MADD	*ETFA*
MADD	*ETFB*
VLCAD	*ACADVL*

B. Disorders considered by different NBS programs or initiatives worldwide	Genes
Cystic Fibrosis	*CFTR*
Congenital Adrenal Hyperplasia	*CYP21A2*
Biotinidase deficiency	*BTD*
3-Methylcrotonyl-CoA Carboxylase	*MCCC2*
Hemoglobin disorders	*HBB*
Hemoglobin disorders	*HBA1*
Hemoglobin disorders	*HBA2*
G6PD deficiency	*G6PD*
Alpha1-Antitrypsin deficiency	*SERPINA1*
Duchenne-Becker dystrophy	*DMD*
Hurler disease	*IDUA*
Hunter disease	*IDS*
Morquio disease	*GALNS*
Maroteaux-Lamy syndrome	*ARSB*
Gaucher disease	*BGBA*
Niemann–Pick A/B disease	*SMPD1*
Pompe disease	*GAA*
Krabbe disease	*GALC*
Fabry disease	*GLA*
X-Adrenoleukodystrophy	*ABCD1*
Severe Combined Immunodeficiency	*IL2RG*
Severe Combined Immunodeficiency	*JAK3*
Severe Combined Immunodeficiency	*IL7RA*
Severe Combined Immunodeficiency	*IL2RA*
Severe Combined Immunodeficiency	*PTPRC*
Severe Combined Immunodeficiency	*CD3D*
Severe Combined Immunodeficiency	*CD3E*
Severe Combined Immunodeficiency	*CD3Z*
Severe Combined Immunodeficiency	*CORO1A*
Severe Combined Immunodeficiency	*RAG1*
Severe Combined Immunodeficiency	*RAG2*
Severe Combined Immunodeficiency	*DCLRE1C*
Severe Combined Immunodeficiency	*PRKDC*
Severe Combined Immunodeficiency	*AK2*
Severe Combined Immunodeficiency	*ADA*
Severe Combined Immunodeficiency	*LIG4*
Severe Combined Immunodeficiency	*NHEJ1*
Severe Combined Immunodeficiency	*CD3G*
Severe Combined Immunodeficiency	*CD8A*
Severe Combined Immunodeficiency	*PNP*
Severe Combined Immunodeficiency	*RMRP*
Severe Combined Immunodeficiency	*ZAP70*
Severe Combined Immunodeficiency	*CD40LG*
Severe Combined Immunodeficiency	*FOXP3*
Severe Combined Immunodeficiency	*IL10RA*
Congenital Hypothyroidism[a]	*TSHR*
Congenital Hypothyroidism[a]	*THRA*
Congenital Hypothyroidism[a]	*THRB*
Congenital Hypothyroidism[a]	*FOXE1*
Congenital Hypothyroidism[a]	*NKX2-1*
Congenital Hypothyroidism[a]	*NKX2-5*
Congenital Hypothyroidism[a]	*PAX8*
Congenital Hypothyroidism[a]	*SLC26A4*
Congenital Hypothyroidism[a]	*FOXI1*
Congenital Hypothyroidism[a]	*KAT6B*
Congenital Hypothyroidism[a]	*KCNJ10*
Congenital Hypothyroidism[a]	*UBR1*
Congenital Hypothyroidism[a]	*GNAS*
Congenital Hypothyroidism[a]	*SLC16A2*
Congenital Hypothyroidism[a]	*TPO*
Congenital Hypothyroidism[a]	*SLC5A5*
Congenital Hypothyroidism[a]	*DUOX2*
Congenital Hypothyroidism[a]	*DUOXA2*
Congenital Hypothyroidism[a]	*IYD*
Congenital Hypothyroidism[a]	*SECISBP2*

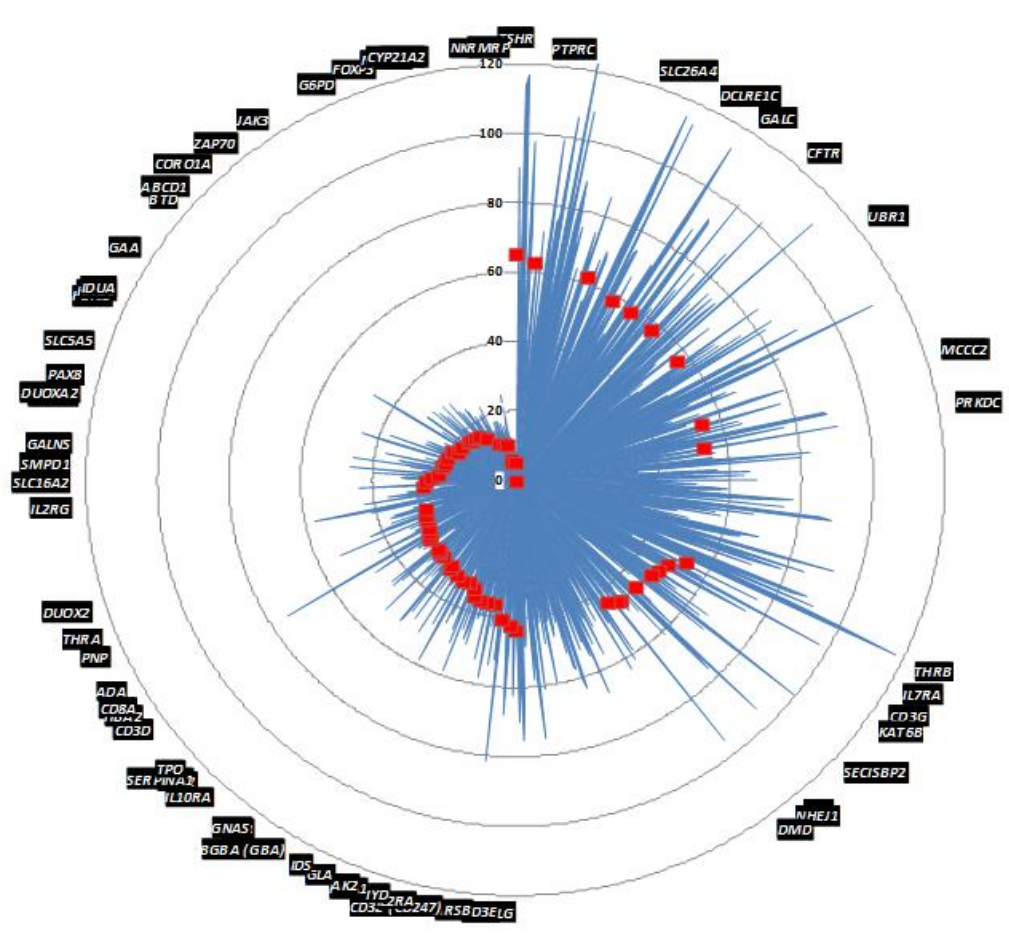
[a] Molecular aetiology of Congenital Hypothyroidism (CH) is not fully understood yet. Only genes currently known as defective in CH are reported.

Coverage depth for the different exons

- 35 IEM genes included in the ONE NBS program



- 65 genes considered by different NBS programs



Identification of known mutations

Patient's Id	Disorder	Gene	Mutation(s) Allele 1	Mutation(s) Allele 2
DBS-1	PKU	*PAH*	c.482T>C	c.1222>T
DBS-2	PKU	*PAH*	c.473G>A	c.842C>T
DBS-3	MCAD	*ACADM*	c.985A>G	c.985A>G
DBS-4	Propionic Acidemia	*PCCB*	c.990_991 insT	c.1252G>A
DBS-5	Methylmalonic Aciduria	*MMAB*	c.556C>T	c.563-577dup
DBS-6	Tyrosinemia type I	*FAH*	c.554-1G>T	c.554-1G>T
DBS-7	Glutaric Aciduria type I	*GCDH*	c.371G>A	c.1204C>T
DBS-8	3-MCC[a]	*MCCC2*	c.1423G>A	c.1535A>C
DBS-9	Propionic Acidemia	*PCCB*	c.997delA	c.763G>A
DBS-10	Homocystinuria	*CBS*	c.429C>G	c.833T>C
DBS-11	PKU	*DHPR*	c.661C>T	c.661C>T
DBS-12	Galactosemia	*GALT*	c.563A>G	c.-119delGTCA c.378-27G>C c.508-24G>A c.507+62G>A c.940A>G
DBS-13	MADD	*ETDFH*	c.293T>A	c.293T>A
DBS-14	MSUD[b]	?[c]	?[c]	?[c]
DBS-15	MCAD	*ACADM*	c.985A>G	c.1091T>C

- All mutations readily detected by WES **but**

- DBS-14 (~10 years old) : MSUD

  Xle = 1262 µmol/L @NBS, AA : allo-Ile => MSUD

  Sanger seq. pending

  WES : no causing mutations in *DBT, BCKDHA, or BCKDHB* genes

  Could not exclude DBT large deletion although coverage is not statistically different from the other 14 samples analyzed

  => Intronic causal mutation ?

- DBS-12 : *GALT*

  4 base-pairs deletion located in the promoter region of Duarte 2 haplotype not covered by our probes

- DBS-5 : MMA

  Identification of 15 base-pairs duplication in *MMAB* gene was critical : not annotated by ANNOVAR, neither automatically identified with IGV soft

"Presumed benign" polymorphisms?

- Number of variants (in 100 genes) annotated in the different samples, and the putative clinical relevance of filtered polymorphism evaluated with 2 different databases

	Variants	Filtered[a] variants	MutationTaster				ClinVar			
			Benign	VUS[b]	Pathogenic	Unknown	Benign	VUS[b]	Pathogenic	Unknown
DBS-1	336	10	0	3	1	6	4	1	2	3
DBS-2	276	9	0	3	1	5	3	0	1	5
DBS-3	317	7	0	1	0	6	3	1	1	2
DBS-4	432	18	0	10	0	8	6	1	1	10
DBS-5	342	9	0	2	1	6	3	0	1	5
DBS-6	339	9	0	3	1[c]	5	5	0	2[c]	2
DBS-7	335	10	0	2	2	6	4	1	2	3
DBS-8	338	10	0	5	1	4	3	2	1	4
DBS-9	429	10	0	3	0	7	3	0	1	6
DBS-10	336	16	1	8	1	6	9	0	2	5
DBS-11	296	6	0	3	0	3	2	1	0	3
DBS-12	350	5	0	0	1	4	2	0	1	2
DBS-13	353	9	0	2	0	7	4	2	0	3
DBS-14	337	8	0	5	0	3	4	2	0	2
DBS-15	337	10	0	3	0	7	3	2	1	4

[a] Filtering criteria's: frequency < 1%, located in exon or splicing site (within the first 8 intronic nucleotides), non-synonymous
[b] Variant of Unknown Significance
[c] 2056C>T nonsense homozygote mutation was identified in DUOX2 gene of DBS-6 patient

"Presumed benign" polymorphisms?

- Number of variants (in 100 genes) annotated in the different samples, and the putative clinical relevance of filtered polymorphism evaluated with 2 different databases

	Variants	Filtered[a] variants	MutationTaster				ClinVar			
			Benign	VUS[b]	Pathogenic	Unknown	Benign	VUS[b]	Pathogenic	Unknown
DBS-1	336	10	0	3	1	6	4	1	2	3
DBS-2	276	9	0	3	1	5	3	0	1	5
DBS-3	317	7	0	1	0	6	3	1	1	2
DBS-4	432	18	0	10	0	8	6	1	1	10
DBS-5	342	9	0	2	1	6	3	0	1	5
DBS-6	339	9	0	3	**1[c]**	5	5	0	**2[c]**	2
DBS-7	335	10	0	2	2	6	4	1	2	3
DBS-8	338	10	0	5	1	4	3	2	1	4
DBS-9	429	10	0	3	0	7	3	0	1	6
DBS-10	336	16	1	8	1	6	9	0	2	5
DBS-11	296	6	0	3	0	3	2	1	0	3
DBS-12	350	5	0	0	1	4	2	0	1	2
DBS-13	353	9	0	2	0	7	4	2	0	3
DBS-14	337	8	0	5	0	3	4	2	0	2
DBS-15	337	10	0	3	0	7	3	2	1	4

[a] Filtering criteria's: frequency < 1%, located in exon or splicing site (within the first 8 intronic nucleotides), non-synonymous
[b] Variant of Unknown Significance
[c] **2056C>T nonsense homozygote mutation was identified in DUOX2 gene of DBS-6 patient**

"Presumed benign" polymorphisms?

- DBS-6 with TYR I : causal homozygote mutation c.554-1G>T in *FAH*
- + Identification of pathogenic nonsense homozygote mutation c.2056C>T (p.Gln686Ter) in *DUOX2* gene (coverage 6x) => known to cause thyroid dyshormonogenesis type 6 and CH.
- 24 years old patient with fully normal thyroid function (repeated normal thyroxin and thyrotropin over years).
- Confirmation on 2nd experiment (coverage 27x) + Sanger sequencing
- How to explain Genotype/Phenotype discrepancy for a truncated protein?

  Mutation located downstream of the Thyroperoxidase active site.

  Variant databases rely on an unique NEJM publication, reporting a single heterozygote patient. No functional studies

  => Should this variant be classified as probably benign??

Analytical conclusions

- Amount and quality of DNA extracted from DBS are adequate to identify pathogenic mutations by high throughput sequencing

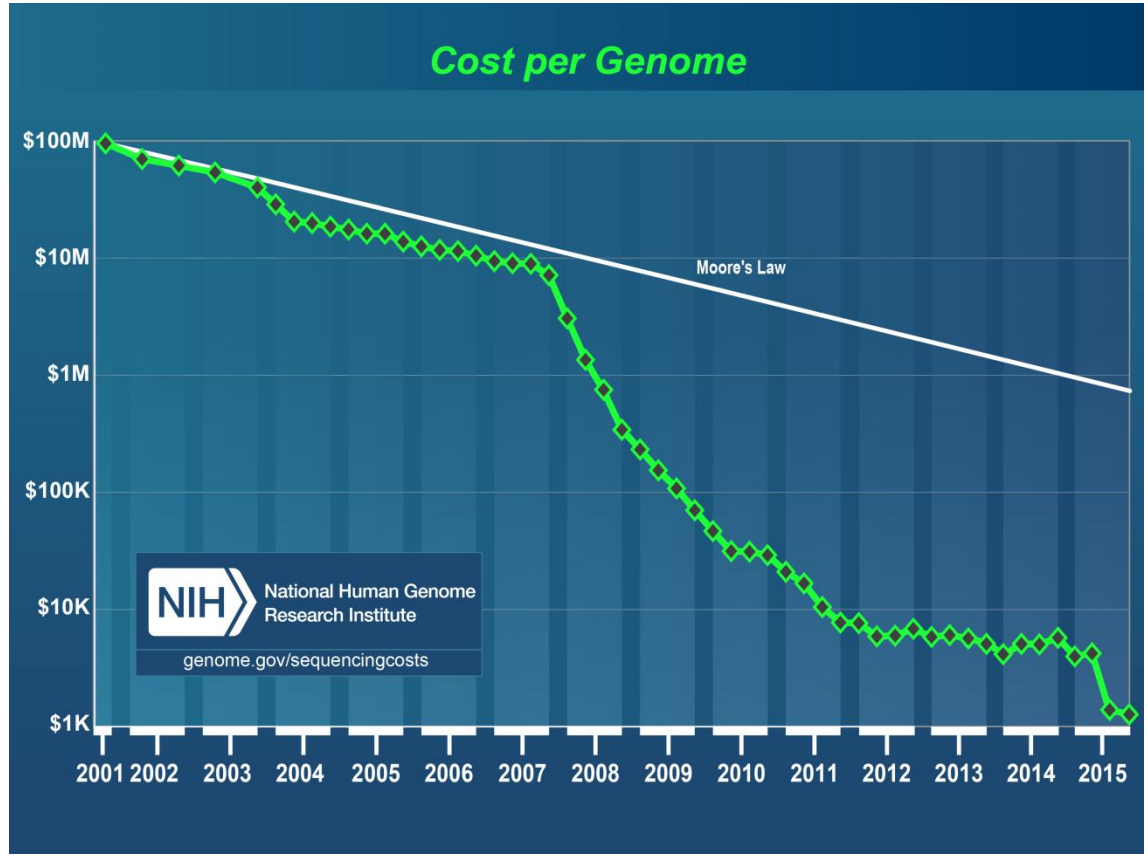
- WES reveals large depth of coverage fluctuations between regions, what could subsequently generate difficulties in variant interpretation.
- Detection of CNV is also subject to caution as unambiguous identification of small or large allelic deletions by NGS can be challenging when the coverage is poor. Moreover, as observed with the 15 base-pairs duplication in the *MMAB* gene, small CNV's are not easily identified by bioinformatics tools.
- Unexpected homozygote pathogen mutation has additionally been characterized in genes unrelated to patient's disorder, questioning the reliability of some variants referenced in databases.

Ready for Next Gen NBS?

- 1st revolution in NBS : introduction of MS 15 years ago.

  2nd revolution underway : NGS as universal approach allowing identification of any disorders with one sole technology.

- Costs



Ready for Next Gen NBS?

- 1st revolution in NBS : introduction of MS 15 years ago.

  2nd revolution underway : NGS as universal approach allowing identification of any disorders with one sole technology.
- Costs
- Poor predictive value (heterozygous carriers, polymorphisms and intermediate deficiencies) => heavy impacts on medical practices and health care budgets.

  Clarification of genotype/phenotype correlations required => Improvement of genomic knowledge and subsequent enhancement of sensitivity and specificity.
- WES generates very large amount of additional data (fastidious bioinformatics treatment and variant's reviewing) => high TAT.

  => Implementation of targeted approach (to easier the coverage homogenization with minimal read depth threshold), apparition of consolidated bioinformatics flowchart

  => Reduction of NGS analysis time.
- List of target genes not to be restrictive, as NBS programs are constantly evolving with new therapies dvpt.
- Earlier blood collection (day of birth, cord blood ?) => anticipation of medical care?

Ready for Next Gen NBS?

- Voluntary expansion of screening providing choice to families who want to know about other conditions?
- Current restriction to diseases with effective treatment ignores potential benefits of any preventive intervention.

  Early identification of patients for other conditions could probably allow assessing pre-symptomatic therapies in randomized studies.
- Huge educational challenges

  Health professionals training and public information

  Parents should be entitled to be informed of the results, of its implications and of the follow-up required.

  Appropriate infrastructures, education, follow-up and psychological assistance should be set up.

  Specific registries to be created.

Ready for Next Gen NBS?

"Next-Generation" NBS is underway

New era questioning the current neonatal screening dogmas

New paradigms have to be agreed

Clinical, political, economical, societal and ethical debates

"Science without conscience is only ruin of the soul"